

THE EPIDEMIOLOGY OF MONITORING<sup>1</sup>Leslie M. Reid<sup>2</sup>

**ABSTRACT:** An informal sample of 30 flawed monitoring projects was examined to identify the most common problems and to determine how they could have been prevented. Problems fall into two general categories: 70 percent of the sampled projects had design problems, and 50 percent of the sampled projects had procedural problems. Monitoring projects implemented by land-management agencies tended to have a higher proportion of procedural problems than did university-based programs (generally graduate student research), while the frequency of design problems was similar between agencies and universities. The most common problems were poorly trained or unmotivated field crews (37 percent of projects, a procedural problem), a sampling plan that was not capable of measuring what was needed to meet project objectives (30 percent, design), delays in analyzing data (27 percent, procedure), inadequate monitoring durations (27 percent, design), and absence of the collateral information needed to interpret results (20 percent, procedure). Most of the problems could have been avoided by submission of the study design to thorough technical and statistical review, active participation of the principal investigators in field data collection, and analysis of at least some of the data as soon as information was collected so that problems could be recognized early enough to be corrected.

(KEY TERMS: monitoring; sedimentation; instrumentation; watershed management; water quality.)

## INTRODUCTION

Monitoring programs are increasingly called for as a way to provide early warning of environmental changes and to evaluate the performance of land management plans so that they can be modified through adaptive management. The Northwest Forest Plan (USDA and USDI, 1994) includes requirements for considerable monitoring work, and the Washington Forest Practices Board considers monitoring to be an important component of watershed

planning (Washington Forest Practices Board, 1995). The California Department of Forestry and Fire Protection is designing monitoring protocols for private timberlands in California, and, at a local level, watershed coalitions are embracing citizen-based monitoring projects as a way to increase public awareness and provide data relevant to local problems. The interest in monitoring has grown markedly over the past decade, and considerable energy, capital, and enthusiasm are now being devoted to collecting monitoring data.

Many texts describe attributes to monitor and how to monitor them (e.g., MacDonald *et al.*, 1991), and there are ongoing efforts to standardize many monitoring protocols. Technological aspects of monitoring are widely researched and discussed. However, observation of several failed monitoring projects suggested that failures resulting from sociological problems may be as common as those resulting from technological ones. Nontechnological problems are rarely considered in monitoring manuals, so it became useful to examine a wider range of failed projects to identify factors contributing to failure.

In addition, scientists at the Redwood Sciences Laboratory (USDA Forest Service, Arcata, California) receive several requests each year to help plan new monitoring programs or to help salvage programs with problems. In the first case, the most common question is a variation of "What type of sampler should we buy?" In the second, most questions are of the form, "We've started analyzing the data, but something seems to be wrong." Thus, on the one hand it appears that practitioners know all they need to about how to design a monitoring project, and all that

<sup>1</sup>Paper No. 98144 of the *Journal of the American Water Resources Association*. Discussions are open until April 1, 2002.

<sup>2</sup>Research Geologist, USDA Forest Service, Pacific Southwest Research Station, Redwood Sciences Laboratory, 1700 Bayview Drive, Arcata, California 95521 (E-Mail: lreid@fs.fed.us).

remains is to find out what brand of pump sampler works the best for suspended sediment. On the other hand, the frequency of "first aid" requests is high enough to suggest that either this level of confidence is unfounded or that many problems cannot be averted by careful project design. Understanding the patterns of failure for monitoring projects would help determine whether problems might best be averted at the design stage or during implementation. Analysis of the reasons for failure would also indicate the specific kinds of problems to be alert for as monitoring projects are designed and implemented.

### SAMPLE SELECTION AND DESCRIPTION

This paper describes the results of an informal survey of 30 monitoring projects that had been diagnosed as being flawed by either participants or reviewers of the projects. The survey was conducted by asking colleagues in land-management agencies, research agencies, universities, and consulting firms to describe their experiences with monitoring projects that had not performed as intended. The sample set thus is not representative of monitoring projects in general, but is intended only to provide an indication of the kinds of problems that might occur. The sample set suffers, in particular, from biases related to human nature (we tend to recall the most spectacular failures) and to sampling of studies from only a few disciplines (primarily hydrology and wildlife biology; other disciplines might characteristically employ other kinds of monitoring strategies). Many of the studies are ones that the informants had not participated in directly but had been asked advice about, but informants were also candid about their own misadventures. Because the intent of the survey was simply to identify the kinds of problems experienced, the specific studies considered here are not identified.

Each study was classified according to the kind of organization responsible for planning and carrying out the work and according to the estimated cost of the project. Nearly one-third of the studies represented cooperative efforts between different organizations. Because the earliest of the surveyed studies had been carried out in the 1970s, costs were estimated according to what the study would have cost today. Only the actual financial outlay was included; the value of borrowed or previously-purchased equipment was not considered, and neither was the value of the time expended by unpaid graduate students. Thus, if the same monitoring project were carried out independently by a graduate student, an agency employee, and a consultant, it might carry three very different estimated costs.

### CLASSIFICATION OF MONITORING PROBLEMS

Each of the 30 studies was first evaluated to identify the kinds of problems that detracted from the project's success. Problems were found to fall into two general categories. Design problems were inherent in the study plan, and no amount of conscientious implementation would have remedied them. Procedural problems, in contrast, were problems in implementation that would derail even well-designed studies. Design problems can only be averted in the planning stage; procedural problems can only be avoided in the implementation phase. Problems were then further classified into the 12 categories listed in Table 1. Most studies exhibited more than one problem.

The most common problem encountered was a procedural problem: 37 percent of the studies were hampered by lack of training or lack of enthusiasm among the people who were to do the work. In one case, the monitoring work was assigned to a field crew located 500 km from the office in which the monitoring plan had been developed. The object of the study was to evaluate the sediment load generated by thunderstorms. As it turned out, most of the thunderstorms that year occurred on Sundays, but the field crew would not work on weekends. In other cases, the field personnel simply did not have the training or initiative to adequately handle unforeseen problems, as was the case with the person who carefully checked the ink supply in a stage-recorder pen every month for a year but did not think it odd that the pen never moved.

Next in importance was a design problem: in nearly one-third of the studies, the sampling plan could not provide the kind of information that was needed to meet the studies' objectives. For example, one study demonstrated unexpectedly that runoff from dirt road surfaces does not contribute sediment to streams. This result was produced because one pump sampler had been positioned upstream of a culvert input and the other was located slightly downstream on the opposite side of the channel. The sediment plume, as one might expect, did not mix quickly enough with the flow to reach the intake of the downstream sampler on the opposite bank.

The third problem was a procedural one that could have averted the effects of other problems if it had not occurred: people did not analyze the data in time to see that there was a problem. The most extreme example of this was a 14-year study of sediment accumulation in weir ponds. The protocol somehow had become confused, so that every year the survey team waited until after equipment operators had emptied the debris basins to resurvey the basins. Predictably,

TABLE 1. Problems Identified in the Surveyed Monitoring Studies.

Nature of Problem		Problem Category	Number	Percent
1.	Nonideal Field Workers	Procedure	11	37
2.	Method Cannot Measure What is Needed	Design	9	30
3.	Data Not Worked Up in Time	Procedure	8	27
4.	Study Too Short	Design	8	27
5.	Collateral Information Missing	Procedure	6	20
6.	Cryptic Technology	Procedure	5	17
7.	Inadequate Problem Analysis	Design	5	17
8.	Fundamental Misunderstanding of System	Design	4	13
9.	Statistically Weak Design	Design	4	13
10.	Personnel Change	Procedure	4	13
11.	Lack of Institutional Commitment	Procedure	3	10
12.	Protocol Changes Prevent Comparison	Procedure	2	7
Total Problems			69	
Total Studies:			30	100
Studies With Design Flaws:			21	70
Studies With Procedural Flaws:			15	50

the volumes of the cleaned-out debris basins remained the same for 14 years. More commonly, this problem appeared as something as simple as uncertainty about exactly which year a benchmark survey stake was moved. In either case, the oversights could have been remedied if data had been analyzed as they became available.

Fourth was another design problem: a quarter of the studies were not of sufficient duration to answer the questions posed. Monitoring programs that were intended to detect management-related hydrologic and geomorphic changes in California during the late 1980s were thwarted by the protracted drought. Even the changes that should have been obvious were not statistically significant because there were too few storms.

About one-fifth of the studies were compromised by the lack of collateral information needed to interpret results. This oversight could be anything from a field-worker neglecting to note whether measurements were in centimeters or inches, to someone trusting the date of a benchmark change to memory instead of notes.

The sixth problem was a procedural one that is particularly compelling in view of the rapidly increasing use of data loggers, geographic information systems, global positioning systems, and other sophisticated tools that are not fully understood by those using them. About 17 percent of the projects

were sabotaged by their technology. In one case, a consulting company used monitoring data from a computer file for more than a year before realizing that the file had somehow become filled with essentially random numbers. In another, it took two years for operators to realize that the hydrologic computer model they were using contained a major programming error that produced results in error by more than an order of magnitude.

The seventh most common problem was again a design problem: 17 percent of the projects diligently collected data that were irrelevant to the problem the projects had been intended to address. The most extreme example of this problem was a series of well-run monitoring studies designed to define values to be used in the Universal Soil Loss Equation (Wischmeier and Smith, 1978) so that reservoir sedimentation rates could be predicted. Studies were finished, the reservoir built, and sediment filled it within a decade. Hindsight revealed that landslides should have been evaluated, not surface erosion. A preliminary problem analysis would have been sufficient to avoid this embarrassment, since landsliding rates could easily have been estimated from aerial photographs.

The eighth problem is related to the previous one: physical and biological systems do not always work the way we think. One monitoring project for an endangered species was discontinued when the principal investigators learned from the results of a new

study that even if declines in population recruitment were occurring now, they would not be detectable for another decade, when it would be too late to respond. In other cases, the preliminary monitoring results themselves provided the information that derailed the project. This situation is not so much a problem as a learning experience. Even though a monitoring program does not meet its original objectives, it still may be successful if it contributes to a better understanding of the system. This problem differs from the previous one because in this case the necessary information is unknowable at the onset of the project. In the previous case, methods for finding the relevant information were well established, and it was simple negligence that prevented its acquisition.

Statistics show up overtly as problem number nine. Thirteen percent of the studies would have been more efficient or more useful if more attention had been paid to their statistical requirements. In one case, redesign of an inefficient sampling strategy would have provided the same statistical power with one-third the number of samples. This problem frequently appears when channel cross sections are monitored. Sample locations are often randomly selected without regard to patterns of change that are already understood. We already know that alluvial channel reaches are responsive to change while bedrock reaches are not, and that pools respond differently than riffles. Stratification of the channel system by characteristics such as these and random selection of sampling locations within the resulting strata would thus represent an efficient sampling strategy. The monitoring programs that did not take this approach showed no statistically significant changes in channel morphology even when changes were readily visible. In one case it was obvious that all of a channel's pools had partially filled with sediment, but only one of the randomly located cross-sections happened to be in a pool.

The tenth problem is related to the first: 13 percent of the studies were affected adversely by personnel changes. In several cases, the original instigator of the project was promoted to an office job, and the people who inherited the field work did not consider the study to be a high priority or did not have the backgrounds required to complete the work. In another case, design and implementation of a monitoring project was contracted out to a consultant. When the contract came up for renewal several years later, the original firm was underbid, and a new consultant took over. When agency personnel attempted to analyze the results several years after that, it became very clear why the second consultant had been able to put in such a low bid.

Other problems that showed up in only a few studies included a lack of institutional support for continuing projects once they had begun, and changes in

protocol part way through a project that made comparison of early and late data impossible.

Many of these problems are related to one another, and most studies had more than one problem. Design problems occurred in 70 percent of the studies, and 50 percent had procedural problems (Table 1). Thus, if all the design problems had been solved, half of the projects would still have failed due to procedural problems. Or if all the procedural problems had been solved, 70 percent would still have failed because of design problems.

## PATTERNS OF FAILURE

Studies were then examined to determine whether there is a pattern to the kinds of studies at risk for particular kinds of problems. The non-random nature of this sample set, its small size, and its inherent biases prevent application of sample-based inferences to the population of monitoring studies as a whole. Results are thus provided merely to allow the strength of patterns within the sample set to be explored.

Studies were first divided into three categories according to their cost: low-cost studies were carried out for less than \$50,000, moderate-cost studies required \$50,000 to \$150,000, and high-cost studies required more than \$150,000. The average cost for the sampled studies is about \$100,000. A chi-square test for differences in proportions (Cochran and Cox, 1959:103) shows a significant trend in the relative frequencies of problems of each type: procedural problems are more common in low-cost studies, while design problems become increasingly common as the cost of the study increases (Figure 1). This pattern is not unexpected: big studies are complicated, so designs are complicated, while shoe-string projects are less likely to carry the commitment that helps avoid procedural problems.

Another pattern appears when studies are examined according to who carried them out. Three categories are distinguished: land-management agencies (including the USDA Forest Service and several foreign ministries); government research agencies; and universities, in which most of the studies represent graduate research work. Nine of the studies involved cooperative efforts between organizations and thus fell into multiple categories; these were considered in each. Consulting studies were excluded because only two were sampled. Studies carried out by the land-management agencies had a higher frequency of procedural problems than did university-based studies (significant at the 0.05 level), while the frequency of design problems was the same for both (Figure 2).

Studies carried out by research agencies held an intermediate position. Again, the pattern is not unexpected. Graduate students, who were responsible for most of the university-based projects, usually carry out their own monitoring. The students' futures depend on a successful outcome, so they are highly motivated to maintain effective procedures and to begin analyzing data as soon as information begins to appear. Studies coordinated by management agencies, on the other hand, are often carried out by people who are not the principal investigators and so are not as able or as motivated to recognize and correct the procedural problems.

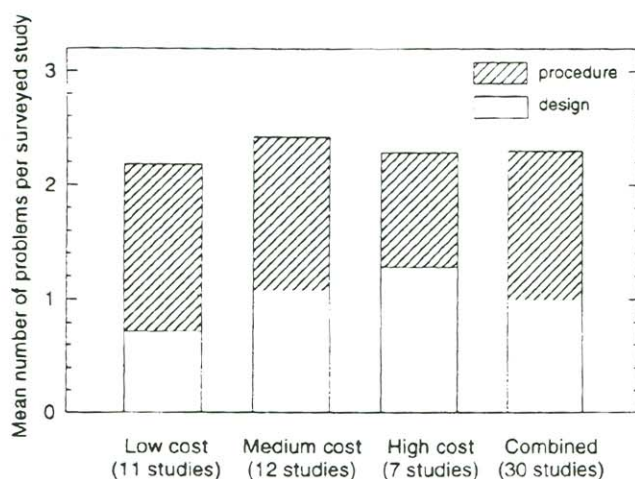


Figure 1. Distribution of Problems by Category for Monitoring Programs Having Different Costs.

While exploring the development of monitoring failures, it became clear that there were distinct patterns in how the problems came about. All of the failed studies incorporated at least one of the following characteristics:

1. A monitoring parameter was selected because it was easy to measure or because it had an established measurement protocol rather than because it was particularly useful for answering the question.
2. The study plan was not subjected to qualified technical or statistical review.
3. The project designers did not have a clear understanding of how the system they were monitoring worked, including what the temporal and spatial scales and magnitudes of response might be.

4. The principal investigators did not directly oversee or participate in the fieldwork.

5. Data analysis did not begin until after the monitoring period ended.

6. The level of commitment, time, or financial support needed to complete the work was underestimated. Often the project budget focused on instrumentation and data collection costs and did not consider the time required for data analysis.

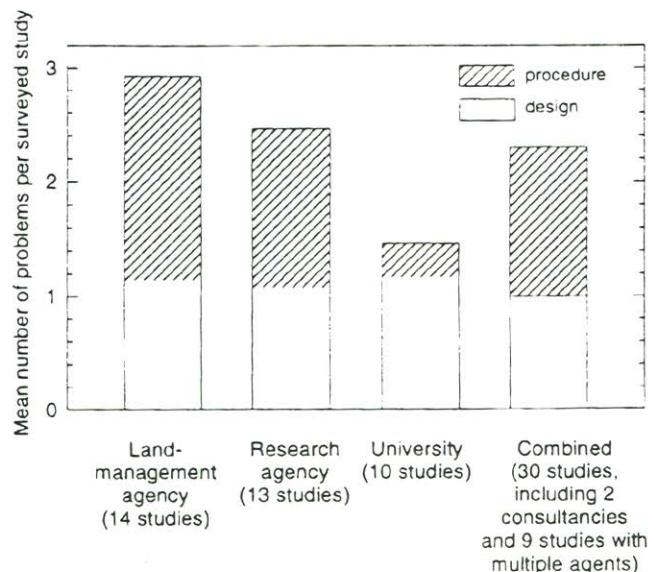


Figure 2. Distribution of Problems by Category for Monitoring Programs Administered by Different Agents.

Each of these characteristics can contribute to the appearance of several of the problems in Table 1. For example, the use of inappropriate monitoring parameters is often a symptom of unclear project objectives, and it is difficult to sustain an enthusiastic field crew in the absence of clear objectives. Unless the project is motivated by a desire to answer a relevant question, monitoring tends to happen for monitoring's sake, and there is little reason to analyze data quickly or to modify the plan to ensure that the question is answerable. Similarly, technical review could alleviate some of the problems caused by a lack of understanding of how the system works and would also help correct underestimates of the effort and resources needed to carry out the project.

## CONCLUSIONS

Given the list of attributes of flawed projects, the measures that might be taken to avoid failure are fairly obvious. First, considerable care must be taken during the design phase to ensure that the planned methods are capable of answering the question being posed. Technical and statistical review at this stage can provide the broad expertise needed to ensure that monitoring parameters are appropriate, that the work can actually be done for the level of funding and effort intended, and that the sampling strategy is statistically sound. Early statistical and technical review is probably the single most effective way to avert design problems.

Procedural problems of all types were reduced in studies where those designing the project were directly involved in gathering the data, and when at least some data analysis occurred as soon as numbers were available. Once monitoring is underway, periodic participation by the principal investigator in the data-gathering process can ensure that the intended protocols are being followed and that relevant complexities are not being overlooked in the interest of consistency. It is also important that data be analyzed as they become available so that procedural oversights can be corrected before it is too late and to be sure that the project is performing as intended.

Although the problems that compromised the surveyed monitoring projects are many and diverse, the existence of these problems should not be interpreted as a rationale for deemphasizing monitoring. Instead, the nature of the problems is, if anything, encouraging – all of those encountered, with the exception of those arising from a fundamental lack of understanding of the system, are easily recognized and can be planned for. In the worst case, early recognition of an insurmountable problem allows the program to be terminated before much work has been done or capital expended. All that is necessary is to know that these problems are common and to take reasonable steps both to prevent them during the design phase of the project and to check for them as the project proceeds. The problems described here only become unmanageable when they are not discovered in time.

## ACKNOWLEDGMENTS

I deeply appreciate the candor of the numerous colleagues who shared aspects of their professional development that I am sure they would have preferred to have forgotten. I also thank two anonymous reviewers for their insights and helpful comments.

## LITERATURE CITED

- Cochran, W. G. and G. M. Cox, 1959. *Experimental Designs*. John Wiley and Sons, New York, New York.
- MacDonald, L. H., A. W. Smart, and R. C. Wissmar, 1991. *Monitoring Guidelines to Evaluate Effects of Forestry Activities on Streams in the Pacific Northwest and Alaska*. U.S. Environmental Protection Agency, Water Division, EPA/910/9-91-001.
- U.S. Department of Agriculture (USDA) and U.S. Department of the Interior (USDI), 1994. *Record of Decision for Amendments to Forest Service and Bureau of Land Management Planning Documents Within the Range of the Northern Spotted Owl; Standards and Guidelines for Management of Habitat for Late-Successional and Old-Growth Forest Related Species Within the Range of the Northern Spotted Owl*. U.S. Government Printing Office, 1994-589-111/00001 Region No. 10.
- Washington Forest Practices Board, 1995. *Standard Methodology for Conducting Watershed Analysis Under Chapter 222-22 WAC (Version 2.0)*. Department of Natural Resources Forest Practices Division, Olympia, Washington.
- Wischmeier, W. H. and D. D. Smith, 1978. *Predicting Rainfall Erosion Losses – A Guide to Conservation Planning*. U.S. Department of Agriculture, Agricultural Handbook 537, 58 pp.